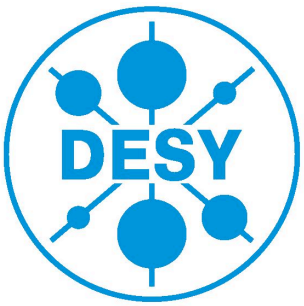

SASE spectrum simulation in CrystFEL

Alexandra Tolstikova, Moscow Institute of Physics and Technology, Russia

September 5, 2013

Abstract

Nonmonochromatic pulses produced by X-ray FEL based on self-amplified spontaneous emission (SASE) were implemented in diffraction simulation program in CrystFEL software suite for serial femtosecond crystallography. Datasets of diffraction patterns with SASE and single-wavelength spectra were generated and analysed.

Contents

1 Introduction

1.1 Serial femtosecond crystallography

Current crystallography methods require mesoscopic crystals that can take many years of research to obtain. A novel concept for structure determination, where single shot diffraction patterns are collected from a stream of nanocrystals, using femtosecond pulses from an X-ray Free Electron Laser (XFEL) was developed a few years ago. Almost half a decade before the availability of any XFEL, researchers had already realized that the very high dose rates delivered by their short and intense pulses might be a suitable strategy for reducing the amount of damage suffered by a specimen during its irradiation, allowing much smaller crystals to be examined than previously thought possible [1]. The validity of this approach was demonstrated in 2005 with a two dimensional specimen using a soft X-ray free-electron laser - FLASH [2], and it has now been applied directly to macromolecular crystallography.
Each crystal is hit by a single X-ray pulse, forming a single diffraction pattern before being vaporized as a nano-plasma burst (Fig. 1). The diffraction patterns are read out by a set of detectors. Fully automated processing using specialized software developed in the Coherent Imaging Division of CFEL at DESY- CrystFEL software suite [4] combines the intensities from all the patterns by a Monte Carlo merging process, leading to a three-dimensional dataset. Such datasets are processed further using standard macromolecular crystallography software to produce electron density maps.

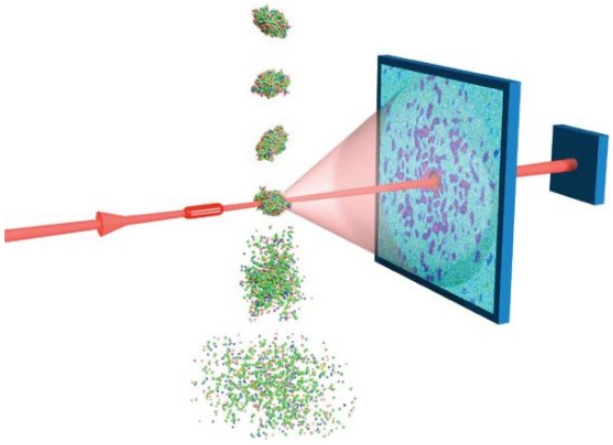

Figure 1: Experimental set-up for atomic resolution images based on single pulse diffraction patterns using the radiation of free-electron laser (FEL) sources such as FLASH in Hamburg (Germany) or LCLS in Stanford (USA).

1.2 Self-amplified spontaneous emission (SASE)

Free-electron laser process is based on self-amplified spontaneous emission (SASE): a relativistic electron pulse from a super-conducting linear accelerator makes a single pass through a periodic magnetic field of an undulator. During the high-gain lasing process, the electrons, perturbed by the magnetic field of the undulator and by their own photon field, form coherent micro-bunches, which behave like a single giant charge, producing strong amplification [2]. Although SASE is very effective, producing tremendously powerful, ultrashort X-ray beams, the start-up from noise leaves poor temporal coherence and a broad, noisy spectrum. A new method, allowing self-seeding using X-rays from the first half of the undulator to seed the second half through a diamond-based monochromator, producing near Fourier-transform-limited X-ray pulses with 0.40.5 eV bandwidth at 89 keV, was presented in recent work [3] (Fig. 2).

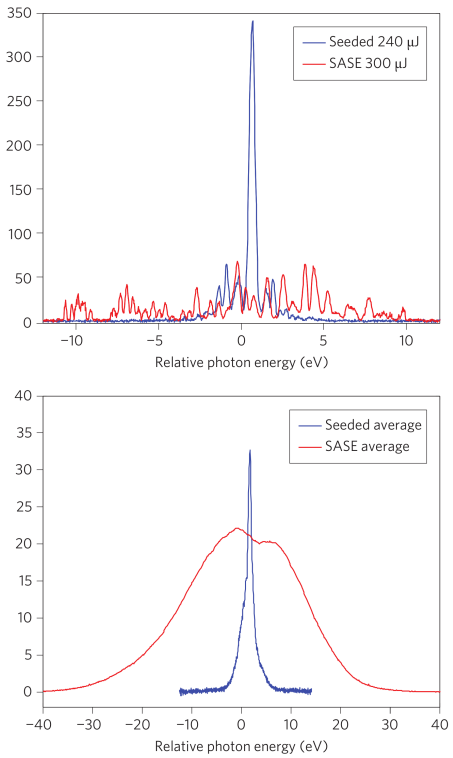

Figure 2: Single-shot (a) and averaged (b) X-ray spectrum in SASE mode (red) and self-seeded mode (blue) [3].

2 SASE pulses diffraction simulation

CrystFEL is a suite of programs for processing diffraction data acquired "serially" in a "snapshot" manner, such as when using the technique of Serial Femtosecond Crystallography (SFX) with a free-electron laser source [4]. It comprises a pattern-sim program for diffraction patterns simulation. The purpose of this project was to introduce SASE spectra to pattern-sim, to generate datasets of diffraction patterns with broad SASE and monochromatic self-seeded spectrum and to analyze and compare them.

2.1 Top-hat spectrum simulation

To simulate broad spectrum one should discretize it, so the first thing to do was to choose the stepsize between samples. I started with simulating top-hat spectra with different width and different number of samples and looking at the diffraction patterns generated (Fig. 3).

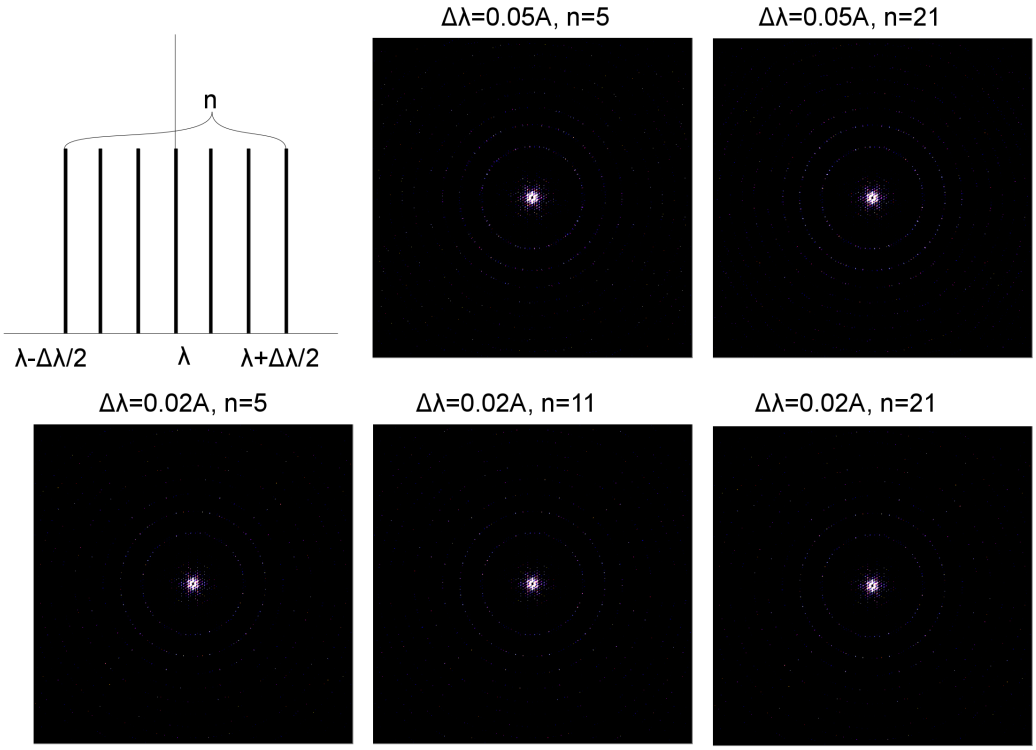

Figure 3: Diffraction patterns generated for different spectrum widths ($\Delta\lambda$) and different number of samples (n)

I tried to reduce stepsize until the pattern stops changing but it appeared to be difficult

to see if there is a difference between patterns just looking at them. So the next action was to write a script which visualizes the difference between two patterns (Fig. 4).

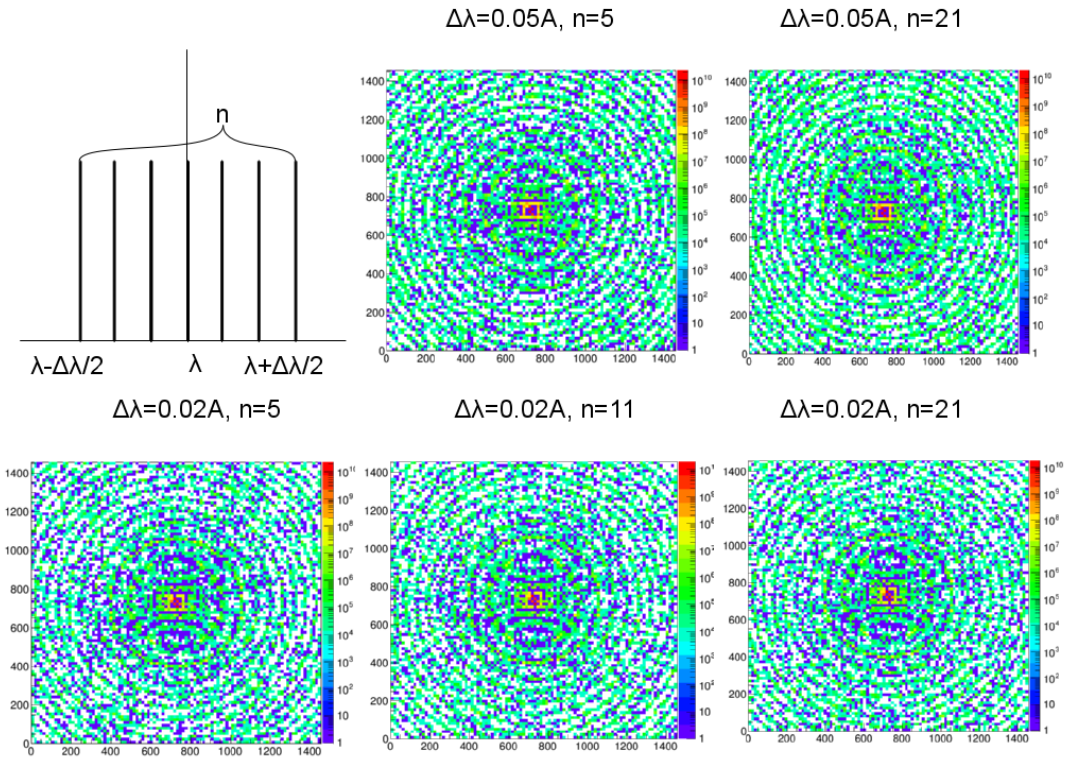

Figure 4: Difference images between diffraction patterns (top-hat - single-wavelength spectra) in the logarithmic scale

Having looked at the difference between top-hat and multiple-wavelength spectra patterns and at the difference images between patterns with different number of samples we came to the conclusion that $0.001\mathring{A}$ stepsize should be enough.

2.2 SASE spectrum generation

Analysis of the dataset of the mesured SASE spectra has shown that it is well approximated by single gaussian with 6eV standard deviation and -7dB gaussian noise (Fig. 5).
Also it was found that the mean pulse energy isn't constant from shot to shot but experiences a random jitter from 6keV in the range of (-13eV, 13eV) (Fig. 7).
We weren't sure if such a significant noise is introduced by the spectrometer or by SASE process. Another problem is that spectra mesured are limited by the spectrometer range

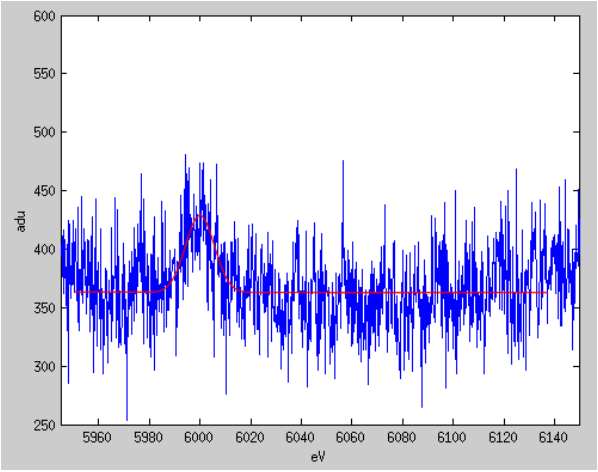

Figure 5: Single-shot SASE spectrum

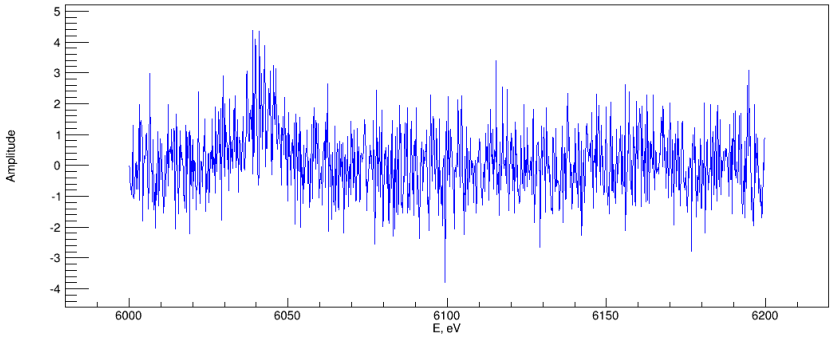

Figure 6: Spectrum simulated in pattern-sim code

$\sim$(5950eV, 6150eV) and we also didn't know what is the resolution of the spectrometer. So the idea was just to simulate such a spectra as we see it at the Fig. 5.
The function generating SASE spectra with the parameters described above was written and introduced to the pattern-sim program code. The result of it's work is shown at the Fig. 6.
Here we came to the problem of sampling again. The mesured spectra had a 0.2 eV sampling stepsize or 1000 samples within the whole spectrometer range. To calculate diffraction from 1000 samples takes too long since the time of pattern simulation is linearly dependent on the number of samples. The idea of sampling with the constant stepsize didn't seem to be good anymore so the next suggestion was to pick some number of highest spikes in spectrum and include them to the diffraction calculation.
To choose this number of samples appeared to be a tricky problem. The quantative figure of merit such as $\Delta = \frac{\sum |X-Y|}{\sum |X|}$ (X - intensities at the 1000-samples diffraction pattern

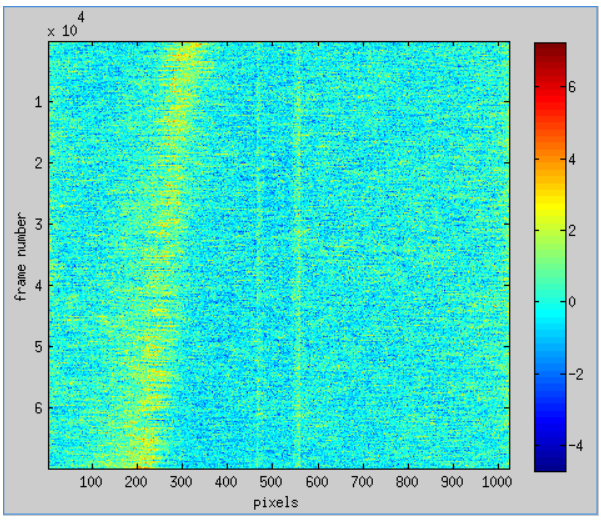

Figure 7: Pulse energy (X-axis) in the set of shots (Y-axis) sorted by jitter value

and Y - intensities at the pattern with reduced number of samples) was introduced. We believed that having calculated this number for different numbers of samples we would derive a criteria, i.e. minimal Δ, when the diffraction patterns become close enough to the 1000-samples spectrum pattern. But unfortunately Δ appeared to increase with the increasing of the number of samples (Table 1) and we still can't explain why.

Number of samples	Δ
1	0,181
20	0,214
50	0,328
100	0,348
200	0,382

Table 1: Δ dependence on number of samples

Having left this problem without further investigation due to the lack of time and having considered how many patterns we want to generate and how much time I have left we've chosen 50 as a number of samples.

2.3 Diffraction patterns simulation

We decided to generate 25000 diffraction patterns with single-wavelength spectra and 25000 patterns with SASE spectra. Pattern-sim includes both c version for calculations using cpu's and opencl version to perform the calculations at gpu's. To speed up the simulations I've modified both versions (they have been a bit different before) to calculate

diffraction of SASE spectrum pulse. Then I've launched the simulations on several gpu and cpu mashines.
Unfortunately when all the simulations had been finished and I started to process the patterns generated I noticed that gpu and cpu versions gave different results in spite of the tests had shown only 0.3% difference between patterns generated by cpu and gpu. Although the mistake in code was eventually found I didn't have time to generate more SASE spectrum patterns. So we only have about 9000 of them.

3 Results and conclusion

3.1 Indexing results

The first step in the analysis of the diffraction patterns is indexing. The results of indexing are presented in the Table 2.
The indexing rate (i.e. number of indexed patterns divided by the total number of patterns) of SASE-spectrum patterns is slightly lower then the indexing rate of single-wavelength patterns. This is not a surprise since the broad spectrum X-ray pulse gives fuzzy Bragg peaks in comparison with a monochromatic one therefore they are more difficult to find and index.
The broad spectrum pulse gives more Bragg peaks because the Ewald sphere corresponding to every sample in the spectrum has different radius hence it crosses different reciprocal lattice points. This was also confirmed by the indexing.
Another very useful figure of merit, perhaps the most useful one in serial femtosecond crystallography, is the self-consistency R-factor, R_{split}. This figure of merit is calculated by splitting the patterns into two, merging each one independently, and then comparing the two resulting sets of intensities [5]. The smaller is R_{split} the better data you have. It has the $R_{split} \sim \sqrt{N}^{-1}$ dependence from number of indexed patterns, so I normalize it multiplying by $\sqrt{\frac{N_{ind}}{25000}}$. The normalized R_{split} of broad spectra patterns is two times smaller then R_{split} of the monochromatic spectra patterns most likely due to the bigger number of peaks per pattern.

	single-wavelength spectrum	SASE spectrum
Total number of patterns	25000	8991
Number of indexed patterns	24916	8944
Indexing rate	99,66	99,48
Peaks per pattern	1315	1770
R_{split}	5,03	4,82
Normalized R_split	5,02	2,82

Table 2: Indexing results

3.2 Conclusion

It can be seen from the results obtained that the broad spectrum pulses give better diffraction patterns. But by the end of this priject we figured out that the results of SASE mesurements (Fig. 5) we tried to simulate in our program were not correct. It appeared that SASE spectrum has much smaller noise especially at the edges of the gaussian, so it needs further investigation.

References

[1] Neutze, R.; Wouts, R.; van der Spoel, D.; Weckert, E.; Hajdu, J., *Nature*, **2000**, *406* (6797), 752-757

[2] Chapman, H. N. *et al.*, *Nature Physics*, **2006**, *2* (12), 839-843

[3] Amann, J. *et al.*, *Nature Photonics*, **2012**, *6*, 693698

[4] White, T. A.; Kirian, R. A.; Martin, A. V.; Aquila, A.; Nass, K.; Barty A.; Chapman H. N., *J. Appl. Cryst.* **2012**, *45*, 335-341

[5] White, T. A., CrystFEL tutorial, http://www.desy.de/ twhite/crystfel/